/// 5 S/-04 N94-27290 00

IMPROVED MODELING OF GPS SELECTIVE AVAILABILITY

P. 10

Michael S. Braasch
Avionics Engineering Center, Ohio University
& Telecommunications and Traffic-Control Systems Group, Delft University of Technology

AnnMarie Fink
Keith Duffus
Avionics Engineering Center, Ohio University

ABSTRACT

Selective Availability (SA) represents the dominant error source for stand-alone users of GPS. Even for DGPS, SA mandates the update rate required for a desired level of accuracy in realtime applications. As has been witnessed in the recent literature, the ability to model this error source is crucial to the proper evaluation of GPS-based systems. A variety of SA models have been proposed to date; however, each has its own shortcomings. Most of these models have been based on limited data sets or data which have been corrupted by additional error sources. This paper presents a comprehensive treatment of the problem. The phenomenon of SA is discussed and a technique is presented whereby both clock and orbit components of SA are identifiable. Extensive SA data sets collected from Block II satellites are presented. System Identification theory then is used to derive a robust model of SA from the data. This theory also allows for the statistical analysis of SA. The stationarity of SA over time and across different satellites is analyzed and its impact on the modeling problem is discussed.

INTRODUCTION

The intentional degradation of the GPS signal known as Selective Availability (SA) is the single largest error source for open loop (non-differential) users of GPS. This degradation is accomplished through manipulation of the broadcast ephemeris data and through dithering of the satellite clock (carrier frequency). Manipulation of the satellite ephemeris data results in erroneous computation of satellite position. This is a long term, non-periodic error trend over the duration of the satellite pass. Dithering of the satellite clock results in erroneous code-phase and carrier-phase measurements. This error trend consists of random oscillations with periods on the order of 5 to 10 minutes.

As the recent literature has shown, a software-centered GPS signal model is essential for the bench testing and evaluation of a variety of GPS-based systems [Bar-Sever, et al, 1990; Braasch, 1990-91; Feit, 1992; Lear, et al, 1992]. A key element in this model is the module for SA. Several SA models have been presented over the past few years; however, each has been derived based on limited data sets or data which have been corrupted by other error sources. An accurate SA-only model is needed. Ideally, this model should be able to generate the typical kinds of SA error traces observed on any satellite at any time. Furthermore, since the two error sources behave quite differently, independent characterization of the orbit and clock components of SA is required. This paper presents work performed to address these issues.

SA DISCUSSION

SA was formally implemented by the Department of Defense on March 25, 1990 [Anon., 1990]. At that time, however, SA had been on experimentally for nearly one year. Various groups reported observing SA-like errors soon after the launch of the first Block II satellite, SVN 14, in February of 1989 [Braasch, 1990-91; Kremer, et al, 1990].

These observations led to the development of the first model of SA based on actual data [Braasch, 1990-91]. In subsequent years, other researchers developed additional SA models [Chou, 1990; Lear, et al, 1992]. None of the investigations, however, were able to answer some fundamental questions: 1) Is SA the same on all satellites? 2) For a given satellite, is SA a stationary random process? That is, do the statistical properties of the SA vary as a function of time? 3) Quantitatively speaking, what is orbital SA?

Presented at the Institute of Navigation 1993 National Technical Meeting, San Francisco, CA, January 20-22, 1993

PRECEDING PAGE BLANK NOT FILMED

PAGE 44 INTENTIONALLY BLANK

ORBIT ERROR ANALYSIS

Accurate modeling of SA requires consideration of both the orbital and clock error components. Previous SA investigations have focussed on the clock component only without consideration of the orbital component.

The ability to observe the orbital error component relies on the data provided by various public and private GPS tracking networks. These networks employ a variety of GPS tracking stations which make range measurements to the satellites. Since the locations of the tracking stations are known, this information can be coupled with the range measurements to calculate the position of the satellites.

The result is the so called precise ephemeris or orbit data. Since the precise orbits are calculated according to where the satellites currently are located, they are more accurate than the broadcast ephemeris data (even without SA) which represents a prediction of where the satellites will be in the future. This precise orbit data is used in a variety of non-realtime GPS applications which require the utmost of accuracy.

The precise orbit data are made available to the public in a variety of formats and media. The data used in this study were obtained from the National Geodetic Survey (NGS) through the Navstar GPS Information Center Bulletin Board and from the Scripps Institution of Oceanography (University of California at San Diego) through their own bulletin board service. The various computer programs required to read the data formats and perform the required interpolations were provided by the NGS [Remondi, 1985; Remondi, 1989; Remondi, 1991]. For verification, precise data were obtained both from NGS and Scripps and compared.

During April of 1992 (days 104, 112, 113), broadcast ephemeris data were collected from 4 Block I satellites and 11 Block II satellites. Some months later, after the precise ephemeris data had been posted, the precise orbits were compared with the orbits calculated using the broadcast ephemeris. Along-track, cross-track and radial errors were calculated and plotted. Since orbit predictions are never perfect, errors on the order of a few meters were expected even in the absence of SA [Ananda, et al, 1984; Bowen, et al, 1985]. Surprisingly, the error plots for all satellites (Block I and Block II) were on the order of a few meters. Figures 1 through 3 show an example of orbital errors computed for satellite 19.

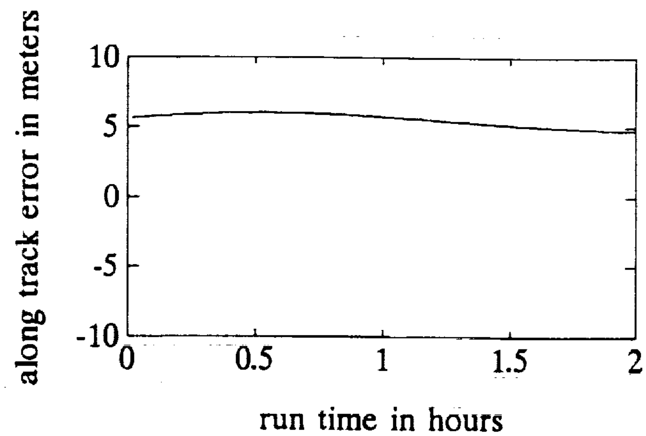

Figure 1. SV 19 ATK Error

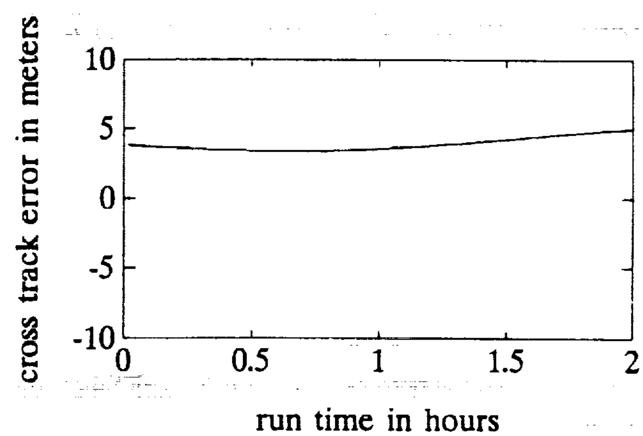

Figure 2. SV 19 XTK Error

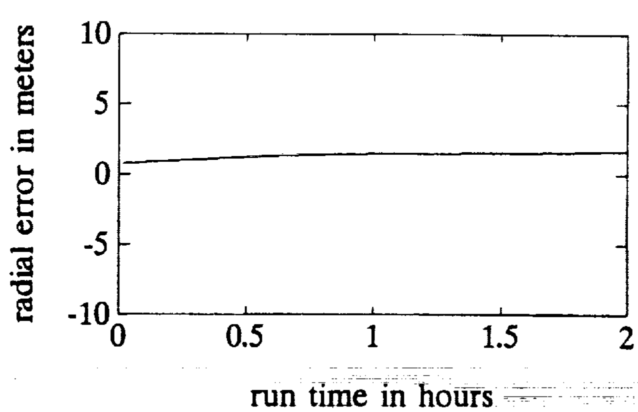

Figure 3. SV 19 Radial Error

Based on these limited data sets, it would seem that the orbital component of SA has not been implemented. It is possible that SA was turned off at this time. However, at the very least, a method now exists whereby the orbital component of SA can be observed. Further data collection efforts are planned to determine if this lack of orbital SA is a regular phenomenon or not.

SA (CLOCK COMPONENT) DATA COLLECTION AND REDUCTION

Having performed the orbital error analysis, the next phase in the study was to collect data for analysis of the clock component of SA. As was noted by Lear, et al (1992), the clock component of SA is a smooth error trace over time and therefore carrier-phase (integrated doppler) data must be collected for the data reduction. This was one of the greatest drawbacks of the models presented in Braasch (1990-91). Since only pseudorange data were available for that study, the data reduction process left a combination of SA and receiver noise. Since filtering could not be performed without imposing assumptions on the underlying SA waveform, it was decided that a model would be derived for the combination of SA and receiver noise [Braasch, 1990-91]. An additional problem with that study was the fact that the data were collected (and hence the model operated) at a data rate of 1/6 Hz. The need for an SA-only model operating at the standard 1 Hz rate served as the original motivation for this study.

During the first week of December (November 30 - December 4), 1992, integrated doppler data were collected at a known location from 10 Block II satellites. The data were collected at Ohio University using a Stanford Telecommunications, Inc. modified Time Transfer System model TTS-502B under the control of a personal computer. The term "modified" refers to the fast-sequencing version of the receiver produced by Stanford Telecommunications, Inc. For the purposes of this study, the important aspect of the modified receiver is its ability to make continuous carrier-phase (integrated doppler) measurements with fine resolution and low noise. The data rate was 1 Hz.

In order to extract the SA waveform, the following steps were taken. First, the true ranges from the satellite to the known antenna location were calculated for the duration of the satellite pass. These were subtracted from the integrated doppler measurements. What remains are referred to as measurement residuals and are a combination of SA, receiver clock drift, atmospheric delay, multipath and a bias due to the ambiguity in the integrated doppler measurements. For environments in which the strength of the multipath is less than the direct signal, the carrier-phase multipath error is guaranteed to be less than 5 cm [Braasch, 1992]. Although it will not be proven here, suffice it to say that the antenna environment used in this study satisfies this criterion. Since a rubidium standard was used as the time base for the receiver, the receiver clock drift is extremely stable and is typically modeled as a first order polynomial [Kremer, et al, 1990]. However, since dual-frequency measurements were not available, ionospheric delay could not be removed. In addition, tropospheric delay is also present. It should be recognized though, that the delays due to the atmosphere are typically long term trends. The result then, is the combination of bias, clock drift and atmospheric delay can be removed by fitting a second-order polynomial to the measurement residuals and subtracting it out. If any bias or long term drift component is present in SA, it will be removed also [Braasch, 1990-91; Lear, et al, 1992]. If an extremely long term error component does exist in the clock SA, it can only be observed if the user clock is synchronized to GPS time [Braasch, 1990-91]. It should also be noted that since the precise ephemerides for the satellites were not available at the time of this writing, broadcast ephemeris was used in the computation of the true ranges. However, under the assumption that the broadcast ephemeris is as accurate as in our previous analysis, this error component is virtually negligible. Even if an orbital SA component is present, it will tend to be removed through the subtraction of the best-fitting second-order polynomial.

The results of the data collection and reduction are shown in figures 4 through 13. The SA error amplitude varies from 40 to 70 meters and the oscillations have periods on the order of 5 to 10 minutes. The variations in the data record length are due to several factors including satellite availability, truncation of records due to receiver glitches and more importantly, truncation of records in order to achieve stationarity. More detail on this last point will be given in a later section.

SA MODEL IDENTIFICATION

Over the past few years, various models have been used to simulate SA. The first SA model was not based on actual SA data but was deduced from a sample probability distribution curve [Matchett, 1985]. The GPS Joint Program Office (JPO) generated SA samples and then computed the curve from these samples. A second-order Gauss-Markov process was postulated and the coefficients were adjusted until its distribution curve matched the one provided by the

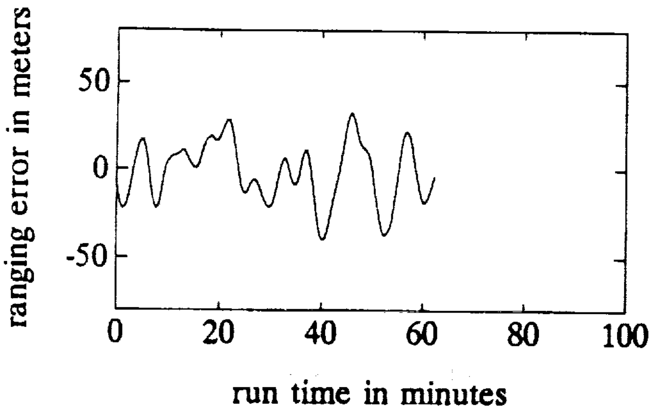

Figure 4. SA on SV 14

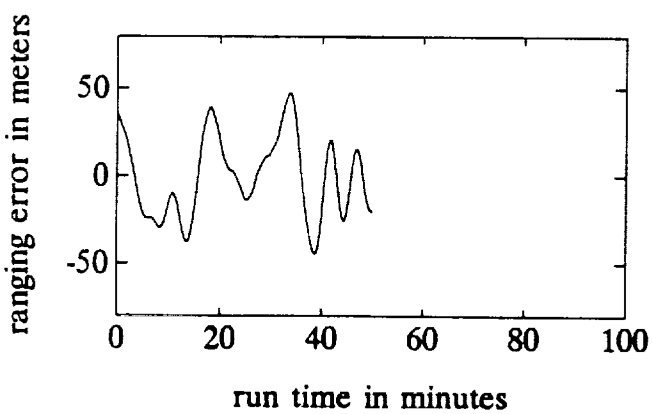

Figure 5. SA on SV 15

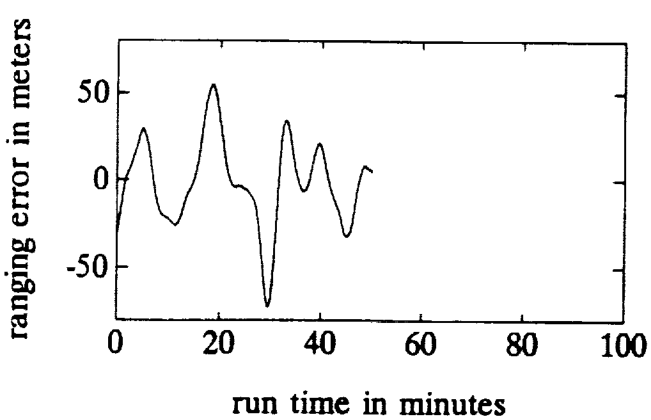

Figure 6. SA on SV 16

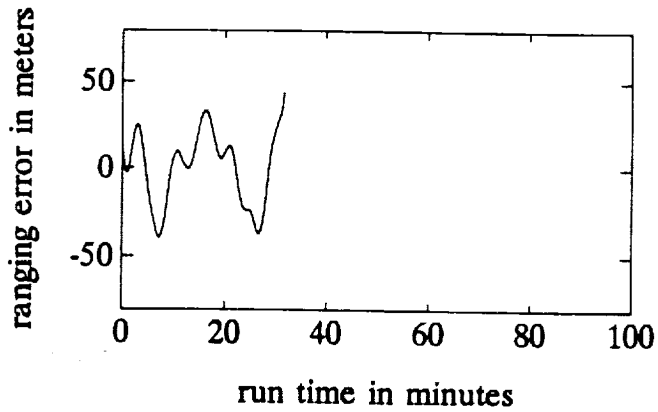

Figure 7. SA on SV 17

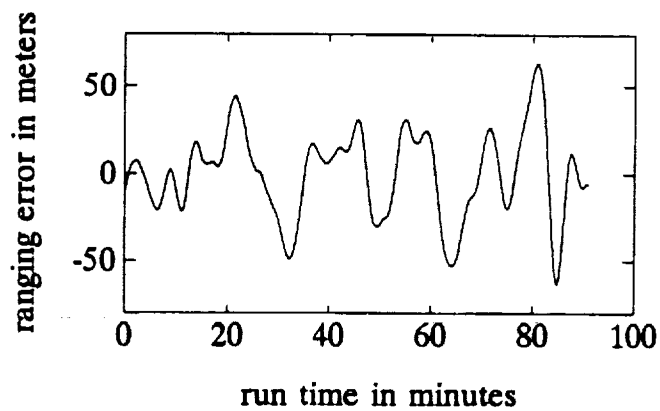

Figure 8. SA on SV 19

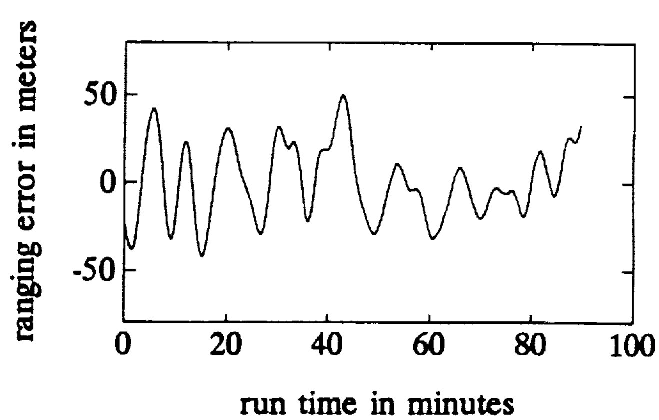

Figure 9. SA on SV 21

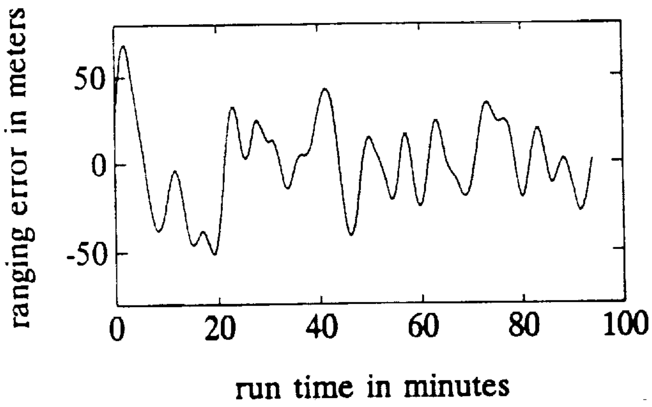

Figure 10. SA on SV 23

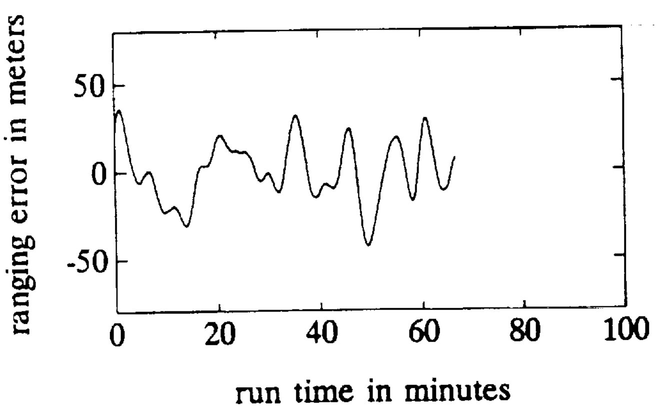

Figure 11. SA on SV 24

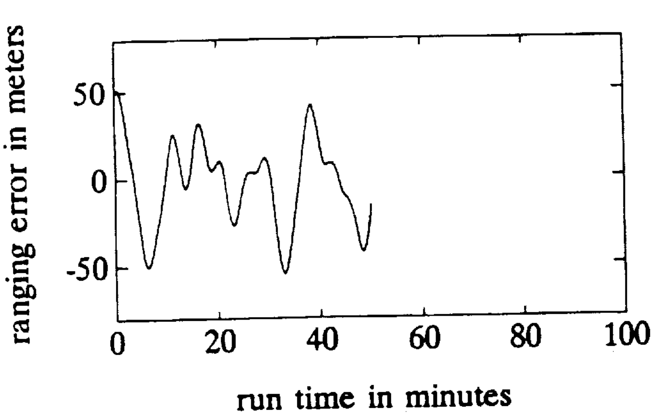

Figure 12. SA on SV 25

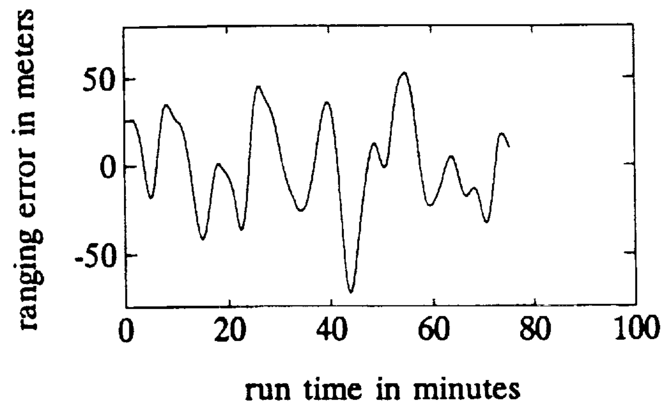

Figure 13. SA on SV 28

JPO. The first models obtained from actual SA data were time series models derived using System Identification theory [Braasch, 1990-91]. Later, Chou also implemented a second order Gauss-Markov process but his was based upon actual SA data [Chou, 1990]. In their recent paper, Lear, et al (1992) present several time series and analytical models also based upon actual SA data.

For this study, System Identification theory was employed to derive time series models in a manner similar to that used in Braasch (1990-91). In general, time series models are based upon the assumption that the data of interest (SA in this case) can be modeled as the output of a linear system (pole-zero filter) driven by Gaussian white noise. Conceptually, derivation of time series SA models can be thought of as a two-step process. The first step is to send the SA data through a filter and adjust the poles and zeros (or equivalently, filter coefficients) such that the output is Gaussian white noise with minimum variance (the output is referred to as residuals). The second step is then to compute the inverse of the filter determined in the first step. Model identification is now complete. Statistically equivalent SA data can then be generated by driving the inverse filter with Gaussian white noise (whose variance is equivalent to that of the residuals in the first step). Kelly (1992) provides an excellent overview of time series model identification and its application to the problem of microwave landing system (MLS) signal modeling.

Three decisions are inherent in the above procedure. The first is the choice of model (filter) type. Three are possible: 1) a pole-zero filter (giving rise to what is known as an Autoregressive Moving Average or ARMA model); 2) an all-pole filter (yielding an Autoregressive or AR model); 3) an all-zero filter (yielding a Moving Average or MA model). The second decision is the choice of model order. That is, if an AR model is chosen, how many poles will be used? The third decision is related to the first two and involves determining if a given residual sequence is white.

Since the primary goal in this study was to derive an accurate SA-only model, an AR model type was chosen. This stems from the fact that ARMA and MA models tend to be noisy. In fact, Braasch (1990-91) concluded that an ARMA model was the best model type for the combination of SA and receiver noise. An autoregressive model of order p (referred to as an AR(p)) is defined as follows [Marple, 1987]:

$$y(n) = -\sum_{k=1}^{p} a(k)y(n-k) + e(n) \qquad (1)$$

where y is the model output, n is the time index, a(k) is the kth filter coefficient, and e is the input Gaussian white noise. Note that the SA models derived from the data will operate at 1 Hz since they are tied to the data collection rate.

Having made the decision to use an AR model type, the rest of the process involved finding the optimum model order and coefficients (pole locations). For a given model order, many methods exist for optimizing the coefficients [Kay, 1987; Ljung, 1987; Marple, 1987]. The one chosen in this study was the Modified Covariance or Forward-Backward method. The second name stems from the fact that the optimization criterion is the minimization of forward and backward prediction errors. As will be shown later, this method performs quite well with SA data.

Several methods exist for model order selection. The majority of these methods have been developed for extremely short data records. The main issue is that one wants to derive a model for the underlying statistical process which gave rise to the data. When model orders are selected which are too high (i.e. approaching the number of data points in the sample), the result is a "fit" of the sample data record rather than the underlying statistical process. The model order selection method used in this study is known as the Principle of Parsimony. The simplest acceptable model is the one chosen. An acceptable model is the inverse of the filter which outputs white noise when driven with SA. Note that if the model order is too low, the residuals will not be white even though the coefficients have been optimized.

The model identification, therefore, proceeds as follows. For a given sample of SA data, the coefficient is optimized for a first-order filter and the residuals are examined. If they are not white, then the coefficients for a second-order filter are optimized and the residuals are examined again. The process is repeated until the model order and optimum coefficients are found for which the residuals are white. This process was performed for each of the SA data sets shown earlier. Depending upon the data set, models of either 9 or 11 coefficients were derived.

The method for determining whiteness involved examination of the autocorrelation function. An example is given in figure 14 where the autocorrelation function is plotted for the residuals from the SA data of satellite 28. Ideally, the autocorrelation function of white noise has a spike at lag 0 and is zero everywhere else. However, that can be obtained only for infinite length sequences. As a result, some minor "sidelobes" will occur at lags other than zero for white noise sequences which are finite. The dotted lines in the figure represent the 99% confidence intervals for the sidelobes. As can be seen in the plot, the sidelobes lie inside the confidence intervals for the most part and thus the model is acceptable.

Further validation of the model can be performed by generating some waveforms and comparing the power spectral densities (PSD's) of the generated and collected data. An example is shown in figures 15 and 16. Figure 15 shows the waveform generated by the SA model which was derived from the SV 28 data. Note that if one compares the waveform to that of the collected data (figure 13), they are not the same. However, they are statistically equivalent. That is, the periods and amplitudes of the generated data are the same as for the collected data. This is better illustrated in figure 16 where the PSD's of the two waveforms are plotted. Although it is difficult to see, there are actually two PSD's plotted. The solid line represents the collected data and the dashed line represents the generated waveform. PSD comparisons were performed on all of the models derived from the data. In each case the result was similar to that shown here.

A final step in model validation concerns the power in the residuals. Recall that in step one of the model derivation process, the goal was to find a filter which output white noise (residuals) with minimum variance when driven with SA. The need for minimum variance is important from both a theoretical and practical viewpoint. Theoretically, having residuals with minimum variance means that the filter has been optimized and embodies the structure (i.e. correlation or information) of the SA. Kelly (1992) refers to this as the filter "explaining" the data. However, from a practical viewpoint, minimum variance is also required. This is particularly true when trying to model random, yet smooth, waveforms such as SA.

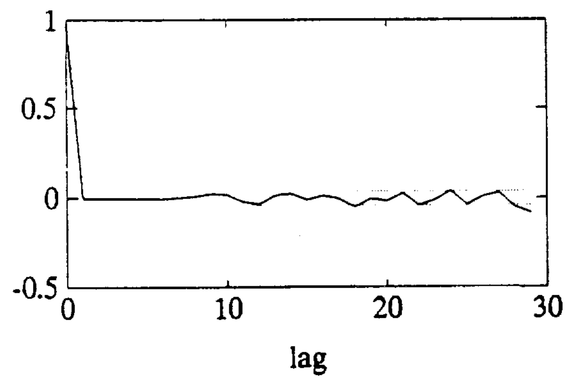

Figure 14. Autocorrelation of SV 28 residuals

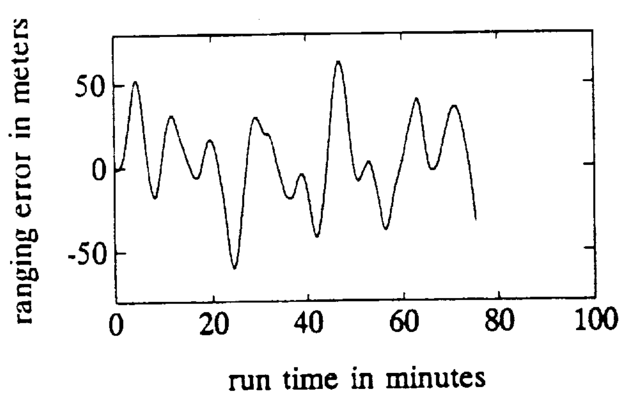

Figure 15. SA model output

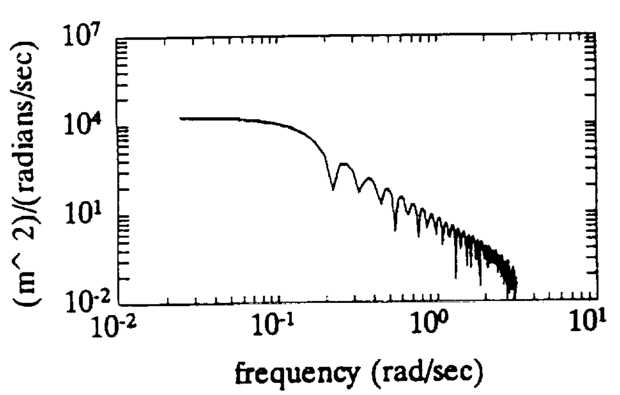

Figure 16. Modeled and measured SA PSD's

Figures 17 and 18 illustrate the success of the AR model type in this respect. The residuals plotted in figure 17 have a standard deviation of 4.12 mm ($4.12\text{x}10^{-3}$ m). Since this represents the amplitude of the noise driving the model (see equation 1), it follows that any noise-like behavior in the generated SA waveforms will be negligible. This is verified in figure 18 which shows the smooth waveform of the generated SA over a short time interval.

MODEL IDENTIFICATION RESULTS

Having derived ten models for SA, the question which poses itself is: Which one do I use? Ideally, one would like to use a single model to generate the SA from all satellites. Multiple SA waveforms corresponding to different satellites could then be generated simply by driving the model with multiple Gaussian white noise sequences. It is therefore necessary to compare the models and the collected data to determine if any equivalence exists. If the collected data share similar PSD's and their corresponding models are similar, then a single SA model is feasible.

As mentioned in the previous section, models with either 9 or 11 coefficients were derived from the collected SA data. For the purposes of comparison, 11th order models were derived for those data sets initially giving rise to 9th order models. Although, strictly speaking, this violates the Principle of Parsimony, the additional complexity of having two more coefficients is negligible.

Although they will not be listed here in their entirety, a comparison of the coefficients for the ten models would seem to indicate little similarity. However, examination of their corresponding pole plots provides more insight. An example is given in figure 19 where the poles of two models are plotted. The models were derived from the data sets of SV 28 and SV 25. The ellipses around the poles indicate the two-sigma confidence regions. Notice that for all of the poles the confidence regions of the two models either overlap or are in close proximity to each other. Admittedly, this is not a strict statistical proof of model equivalence (for that, a multivariate analysis of variance hypothesis test is required; see Kelly (1992)). However, it is at least an indication of model similarity.

Pole-plot comparisons were performed with all of the models. Five were found to be similar. These five were the models derived from SV's 28, 25, 19, 16 and 15. The similarity was verified through comparison of the PSD's of collected and generated SA waveforms. Since the five models are similar, any one of them can be chosen and used as the SA model. The coefficients

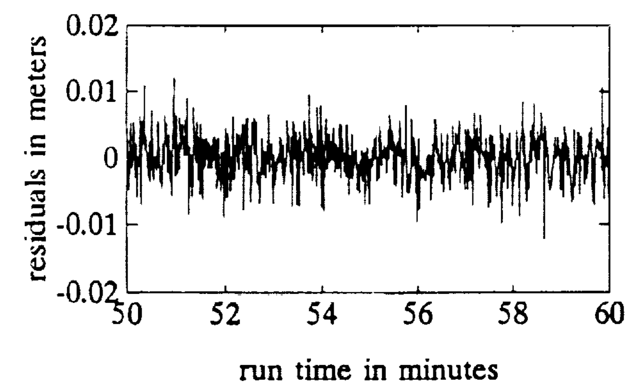

Figure 17. SV 28 residuals

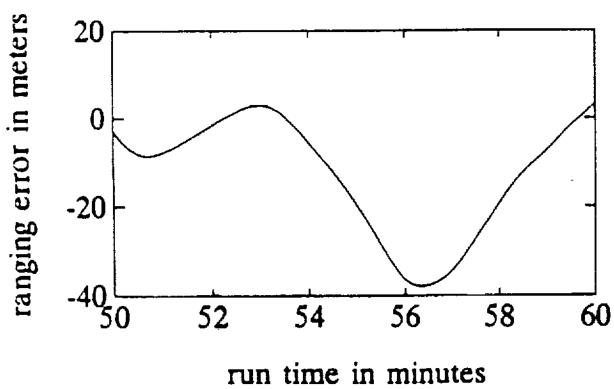

Figure 18. SA model output - expanded scale

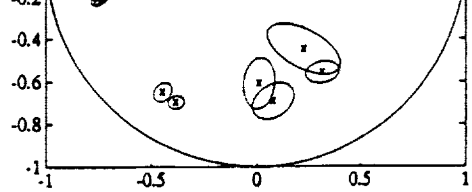

Figure 19. Pole comparison

for the model derived from the SV 28 data will be listed here:

a(1) = -1.36192741558063
a(2) = -0.15866710938728
a(3) = +0.13545921610672
a(4) = +0.21501267664869
a(5) = +0.30061078095966
a(6) = -0.12390183286070
a(7) = +0.10063573000351
a(8) = +0.02694677520401
a(9) = -0.12898590228866
a(10) = +0.05083106570666
a(11) = -0.05600186282898

$\sigma_e^2 = 1.6993 \times 10^{-5}$ (meters^2)

σ_e^2 is the variance of the Gaussian white noise input. The seemingly excessive amount of significant figures is required to ensure filter stability. Note in figure 19 that three out of the eleven poles are extremely close to the unit circle. Truncation of the coefficients can cause these poles to move outside the unit circle yielding instability. It is thus very important that the significant figures be maintained. Towards this end, it is suggested that double-precision arithmetic be employed in the generation of SA waveforms using this model.

The distribution of these poles makes sense from the point of view of filter theory. The three poles grouped near the unit circle and the real axis represent a type of low-pass filter with an extremely narrow bandwidth. This is necessary since the input to the filter is wide-band noise and the output is extremely narrow-band SA. Although the low frequency components dominate the SA waveform, higher frequencies are present also and the other poles of the model contribute to these components.

Stationarity

As was mentioned earlier, some of the collected data records had to be truncated in order to achieve stationarity. A random process is said to be stationary if its statistics do not change with time. Unfortunately, some of the original collected SA records did exhibit non-stationary behavior. Another way of viewing this is to assume that SA is truly generated by a time series model but that the coefficients change as a function of time. Powerful as they are, the vast majority of model identification techniques assume a stationary data record. Non-stationary records are typically examined by segmenting the data into stationary sections and identifying a model for each one separately. The non-stationary behavior of the data then can be determined by examining the change in the models from segment to segment [Marple, 1987].

Having SA with this kind of behavior makes sense, at least from a security point of view. A non-stationary random process is much harder to "crack" than a stationary one. It should be pointed out, however, that the collected data did exhibit stationarity for periods of up to one and a half hours. Since this was the maximum data collection period, no conclusions can be made for longer periods. Future data collection efforts are being planned to examine this phenomenon more closely. In the mean time, the SA models derived from the data are good approximations to the truth.

CONCLUSIONS AND RECOMMENDATIONS

Simulations are often necessary in the process of development and testing of GPS-based systems. For those users of GPS not having the benefits of DGPS corrections, SA represents the dominant source of error. For would-be developers of DGPS systems, SA dictates the trade-off between the update rate (of the differential corrections) and system accuracy. Simulations therefore must account for SA. In this paper, the issue of SA analysis and modeling has been revisited. Using post-processed, precise ephemeris data, a technique has been described whereby the clock and orbital components of SA can be identified separately. For the data collected for this paper, the orbital component of SA seems not to have been implemented.

SA data (clock component) has been collected from over half of the current Block II satellites and a robust model has been derived. The model has been demonstrated to be accurate and robust. It is suggested that this model be implemented in GPS receiver test equipment and in GPS-based system simulations. Since the model is capable of generating virtually unlimited amounts of data, the design and test engineers need not be constrained to a few collected waveforms.

ACKNOWLEDGEMENTS

The work presented in this paper was funded in part by the National Aeronautics and Space Administration under the Joint University Program for Air Transportation Research Grant NGR 36-009-017. Ashtech, Inc. is thanked for their provision of the GPS receiver which was used in the collection of the broadcast ephemeris data. Dr. Frank van Graas of the Avionics Engineering Center is thanked for his support.

REFERENCES

Ananda, M., H. Bernstein, R. Bruce, K. Cunningham, W. Feess, P. Jorgensen, M. Menn and C. Price (1984) "Autonomous Navigation of the Global Positioning System Satellites," AIAA Guidance and Control Conference, Seattle, WA, August.

Anonymous (1990), "Newsfront" GPS World, May/June.

Bar-Sever, Y., T. Yunck and S. Wu (1990), "GPS-Based Orbit Determination and Point Positioning Under Selective Availability," ION GPS-90, Colorado Springs, CO, September 19-21.

Bowen, R., P. Swanson, F. Winn, N. Rhodus and W. Feess (1985), "Global Positioning System Operational Control System Accuracies," NAVIGATION: Journal of the Institute of Navigation, Vol. 32, No. 2, Summer.

Braasch, M. (1990-91), "A Signal Model for GPS," NAVIGATION: Journal of the Institute of Navigation, Vol. 37, No. 4, Winter.

Braasch, M. (1992), "On The Characterization Of Multipath Errors In Satellite-Based Precision Approach and Landing Systems," Ph.D. dissertation, Department of Electrical and Computer Engineering, Ohio University, Athens, Ohio, June.

Chou, H. (1990), "An Anti-SA Filter for Non-differential GPS Users," ION GPS-90, Colorado Springs, CO, September 19-21.

Feit, C. (1992), "GPS Range Updates in an Automatic Flight Inspection System: Simulation, Static and Flight Test Results," ION GPS-92, Albuquerque, NM, September 16-18.

Kay, S. (1987), Modern Spectral Estimation: theory and application, Prentice-Hall, Englewood Cliffs, NJ.

Kelly, R. (1992), "MLS System Error Model Identification and Synthesis," IEEE Trans. on Aerospace and Electronics Systems, Vol. 28, No. 1, January.

Kremer, G., R. Kalafus, P. Loomis and J. Reynolds (1990), "The Effect of Selective Availability on Differential GPS Corrections," NAVIGATION: Journal of the Institute of Navigation, Vol. 37, No. 1, Spring.

Lear, W., M. Montez, L. Rater and L. Zyla (1992), "The Effect of Selective Availability on Orbit Space Vehicles Equipped with SPS GPS Receivers," ION GPS-92, Albuquerque, NM, September 16-18.

Ljung, L. (1987), System Identification - Theory for the User, Prentice-Hall, Englewood Cliffs, NJ.

Marple, S. L. (Jr.) (1987), Digital Spectral Analysis with Applications, Prentice-Hall, Englewood Cliffs, NJ.

Matchett, G. (1985), "Stochastic Simulation of GPS Selective Availability Errors," Technical Memorandum, FAA Contract DTRS-57-83-C-00077, June.

Remondi, B. (1985), "Distribution of Global Positioning System Ephemerides by the National Geodetic Survey," First Conference on Civil Applications of GPS - ION, Greater Philadelphia Section, Warminster, PA, September 12.

Remondi, B. (1989), "Extending the National Geodetic Survey Standard GPS Orbit Formats," NOAA Technical Report NOS 133 NGS 46, National Geodetic Information Branch, National Oceanic and Atmospheric Administration, Rockville, MD, November.

Remondi, B. (1991), "NGS Second Generation ASCII and Binary Orbit Formats and Associated Interpolation Studies," Twentieth General Assembly of the International Union of Geodesy and Geophysics, Vienna, Austria, August 11-24.